DOCUMENT RESUME

ED 207 277 EC 140 011

AUTHOR Bricker, Diane; And Others
TITLE Early Intervention: A Plan for Evaluating Program Impact.
INSTITUTION California Univ., Los Angeles.; Oregon Univ., Eugene. Center on Human Development.; Washington Univ., Seattle. Western States Technical Assistance Resource.
SPONS AGENCY Office of Special Education (ED), Washington, D.C.
REPORT NO WESTAR-Ser-10
PUB DATE May 81
CONTRACT 300-77-0508
GRANT G007701817; G0077001904
NOTE 41p.
AVAILABLE FROM WESTAR, Western States Technical Assistance Resource, 345 N. Monmouth Ave., Monmouth, OR 97361.

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Data Collection; *Disabilities; Early Childhood Education; *Evaluation Methods; *Intervention; Models; *Program Evaluation; Records (Forms)

ABSTRACT

The paper describes the major evaluation problems facing early intervention projects and the solutions for these problems adopted by the Preschool Program, Center on Human Development, which serves handicapped and nonhandicapped children (birth to 5 years) and their families at the University of Oregon. Five components of the program are briefly outlined: intervention in the classroom and in the home, parent involvement, support services, training and dissemination, and evaluation. Difficulties in evaluation are considered for four categories: resources for evaluation, framework for guiding evaluation efforts, suitability of existing measures, and appropriateness of available analytic designs. Among the solutions offered are the following: whenever possible, collect data which have multiple purposes; after identifying the goals of the program, select instruments that reflect the program's emphasis; when no standard instruments are available that match program goals, define in clear behavioral terms what those goals are, then use those definitions as the basis for developing an assessment measure; and evaluate child progress using multiple indices and conditions when possible. Outlined is a plan for documenting child progress in the Preschool Program at the Center on Human Development. A plan is outlined for documenting child progress in terms of resources, framework, assessment measures, data analyses, data preparation, simple data description, and estimation of program effects. Tables and figures provide information on instruments used in the evaluation plan, testing schedule, and training targets. Among the tables and figures provided are a description of instruments used in the evaluation plan, sample recording forms, and flow charts of important considerations in instrument selection, data preparation, and data analysis. Appended materials include descriptions of assessment instruments, a sample demographic form, and additional analytic strategies for assessing program impact. (SB)

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

ED207277

EARLY INTERVENTION:

A PLAN FOR EVALUATING PROGRAM IMPACT[1]

Diane Bricker, Ph. D.
Center on Human Development
University of Oregon

Robert Sheehan, Ph. D.
Purdue University

David Littman, Ph. D.
Center on Human Development
University of Oregon

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Marcia J. May
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

[1]Preparation of this paper has been supported in part by Grant No. G007701817 to the Center on Human Development, University of Oregon, from the Office of Special Education, USOE and in part by Grant No. G0077001904 to Project REACH, University of California at Los Angeles from the Office of Special Education, USOE.

The authors would like to thank Russell Gersten, Laurel Carlson and Bill White for their critical reading of the paper, Gary Harrison and Susan Edginton for their helpful editing, and Connie Nosbisch for the preparation of the manuscript.

WESTAR Series Paper #10

EC 140011

WESTAR Coordinator:

Meave Stevens

Editing by

Gary Harrison and
Arnold Waldstein

Word Processing by

Esther Sugai

Prepared and distributed by:

WESTAR
Western States Technical Assistance Resource
345 N. Monmouth Avenue
Monmouth, Oregon 97361
503/838-1220, ext. 391

Reproduced and distributed under contract number 300-77-0508 from the Office of Special Education, U.S. Department of Education. The opinions expressed herein do not necessarily reflect the position or policy of the U.S. Department of Education, and no official endorsement should be inferred.

Printed in the United States of America

May 1981

Introduction

Educational programs and allied health services for handicapped infants and preschool children have grown increasingly available during the past decade (Bricker, Seibert & Scott, in press). In large measure, this growth in services is due to federal legislation such as P. L. 94-142, which provides states with modest incentives for serving preschool handicapped children until 19..., and to state regulations establishing educational services for young handicapped children. Cohen, Semmes and Guralnick (1979) report that by 1979, 31 states had some form of mandated special education services for children under age six.

Although federal and state legislation guarantees young handicapped children the right to early education, continued support from both federal and state governments is dependent in part upon the ability of early childhood special educators to demonstrate that early intervention has significant educational impact upon young handicapped children. If special educators cannot objectively demonstrate that early intervention produces long-range, beneficial differences in handicapped children's educational achievements, public support for early intervention programs stands to be severely reduced or completely eliminated. If special educators, however, can show that early intervention programs produce a significant facilitative impact on young handicapped children, public support is likely to continue, at least at marginally sufficient levels.

Lack of objective support for the effectiveness of various educational strategies being used with young handicapped children is a serious problem (Bricker, 1978; Clarke & Clarke, 1976). The lack of empirical verification of program impact can be attributed to several factors, including inadequate assessment instruments, inappropriate designs and impractical documentation methodologies (Sheehan, 1979, 1980). The absence of a framework to guide the documentation of child progress further complicates the problem of objective demonstrations of effectiveness. Thus, an urgent need exists for appropriate and practical evaluation strategies. This paper is directed toward this need and describes an early intervention program's comprehensive evaluation plan which was focused on documenting child progress. While we recognize that a complete evaluation plan should also address not only such areas as child progress, but parent involvement, staff training and other important activities, evaluation of these areas is unfortunately less refined than strategies for monitoring child change and therefore remains a major challenge for the eighties.

This paper describes the major evaluation problems facing early intervention projects and the solutions for these problems adopted by the Preschool Program, Center on Human Development, at the University of Oregon, Eugene. Rather than focusing on the quantitative results of the program, we attend primarily to the manner in which the program addressed the issues raised in the discussion of evaluation problems. We hope that this approach will provide a useful model for other intervention programs facing the task of evaluating program impact.

Program Description

Because of the diversity of its target population and the congruence of its programming with other early childhood intervention programs, we have selected the Preschool Program at the University of Oregon as a useful model from which to discuss evaluation. The Preschool Program served handicapped and nonhandicapped children ranging in age from birth to five years and their families; the children's handicapping conditions evolved from numerous etiologies and ranged from mild to profound. Although such diversity among the target population tends to complicate the problems of documenting child progress, the evaluation

strategies suited for such a diverse population are more universally applicable than those suited only for a more specific population.

Like many early childhood intervention programs, the Preschool Program was composed of five separate but integrated components: intervention, parent involvement, support services, training/dissemination and evaluation. Each component is briefly described here to place the evaluation plan in perspective.[1]

The intervention component, divided into a classroom unit and a home intervention unit, afforded classroom and home instruction in gross motor, fine motor, sensorimotor, social, self-help and communication. Instruction at the Center and in one public school classroom was conducted primarily in small and large groups, although individual instruction was offered when necessary. The basic goal of classroom intervention was to develop and implement for each child the most effective educational program, and each child engaged in 15 to 20 instructional activities interspersed with periods of exploratory play daily. Through home intervention, the staff trained parents to become effective teachers of their children and made weekly home visits to observe progress and model new parent activities. The range in age of children in both units was 4 months to 5 years; their handicapping conditions ranged from mild to severe. Several at-risk and nonhandicapped children participated in the classroom unit.

The emphasis of the parent involvement component was threefold: educational, social service and advocacy. In addition to individual and small-group instruction, the staff arranged weekly and monthly parent meetings and coordinated social service counseling for the parents based on their needs. Because the staff encouraged parents to become articulate spokespersons not only for their own child but for preschool handicapped children in general, they supported and informed the parents about advocacy activities.

A physical therapist, psychologist, speech pathologist, social worker and feeding specialist were involved in the support service component. Because this program was unable to support the cadre of needed professionals on a full-time basis, the program adopted a model called "the educational synthesizer" (Bricker, 1976). In this model, the specialist functions primarily as an evaluator and consultant who subsequently monitors the implementation of the developed program. In such a model, the classroom staff and parents rather than the specialist provide the primary hands-on therapy to the child. In this approach, the teacher and parent are responsible for organizing ideas from other disciplines into an integrated developmentally sound approach.

The training and dissemination component covered all aspects of the program including evaluation, curriculum and administration. Graduate students from a variety of disciplines used the program as a practicum; local, state, national and international visitors came to the program site for demonstrations of early intervention procedures; and the program generated data and information for writers of articles and reports.

As we shall fully explore below, the primary objective of the evaluation component was to document child progress. Assessment procedures fell into four categories: the collection of demographic information on the children and families within the program, the pre and post administration of standardized or norm-referenced tests, the pre and post administration of criterion-referenced tests and the collection of daily and weekly data on progress toward IEP target objectives. The remainder of this paper will discuss this component in detail.

---

[1]For detailed descriptions of project components see: Bricker, D. A handicapped children's early education program: Rationale, program description and impact. Final report for the Division of Innovation and Development, Office of Special Education, January 1981.

Evaluation Problems Facing Early Interventionists

Lack of objective evaluation of the effects of early intervention with handicapped infants and preschoolers does not result from indifference or neglect on the part of the interventionists. Evaluation of program effects is an arduous and time-consuming undertaking laced with many complexities and problems. These difficulties can be divided into four distinct categories of problems:

1. Resources for evaluatior
2. Framework for guiding evaluation efforts
3. Suitability of existing measures
4. Appropriateness of available analytic designs

In the following sections, we will look closely at each of these categories and offer possible solutions to the evaluation problems inherent in each.

Resources for Evaluation

A serious problem often overlooked by critics of early intervention programs is the lack of necessary resources to pursue evaluation activities. A sound evaluation plan often requires the assistance of diagnosticians and psychometricians, individuals with research design, statistical and computer analysis skills, and funds for test protocols, computer scoring and analysis time. In addition to the need for tangible support services, there is a need to maintain necessary commitment of project personnel to the evaluation of program and/or child progress. Successful implementation of any evaluation approach will depend upon an entire staff's willingness to support data collection efforts.

Although there are no completely satisfactory solutions to the problems encountered when attempting to assemble and maintain needed support services, some potentially useful approaches to acquiring minimal assistance are suggested below:

1. Contact local colleges or universities to identify persons with appropriate training. They may be interested in using an intervention program as a practicum site in return for consultation. In addition, graduate students with evaluation expertise may be available to assist the program.
2. Determine whether funds can be specifically identified for a part-time evaluator, or if an evaluation specialist can be shared with another intervention program.
3. If they are available, use practicum students involved in the intervention program who have access to computer time for a class project dealing with the program.
4. Whenever possible, collect data which have multiple purposes. For example, the Uniform Performance Assessment System (White, Edgar & Haring, 1978) is used at the University of Oregon Preschool for programming as well as for documenting child progress.
5. If possible, select at least some evaluation instruments or procedures that can be administered by classroom personnel. Instruments such as the Uniform Performance Assessment System (UPAS) and the Developmental Profile (Alpern & Boll, 1972) were designed for use by teachers.

Thoughtful application of strategies such as those suggested above should allow most programs to maximize the use of resources available to collect information pertaining to child progress.

Framework to Guide the Evaluation Effort

A sound and effective educational program needs to be governed or directed by some broad set of guidelines or framework. Such a framework should provide cohesiveness and consistency to a program by directing decision-making in a number of areas including: 1) the determination of short-term and long-term objectives and priority areas for the target child's individual education program; 2) the selection of strategies for facilitating attainment of the established objectives; 3) the selection of appropriate evaluation instruments to assess initial levels of entry into the program and to monitor subsequent change; and 4) the construction, adaptation and/or modification of training materials and curricula.

The adoption of a framework will also assist in the intelligent selection of specific instruments and subsequent analyses to be performed on the data yielded by the chosen instruments. The need for an organized evaluation plan becomes clear when one considers the implications of following "hunches" or conducting analyses because they "might show something." To avoid collecting a morass of data that do not lend themselves to the objectives of the program, we suggest that attention be given to the following strategies.

1. After identifying the goals of the program, select instruments that reflect the program's emphasis.
2. Identify which of the selected instruments or subtests of instruments most closely approximates the goals of the program's intervention. Also identify which of the selected instruments is the most adequate psychometrically.[2] These may not be the same. If the instruments differ, attempt to relate the program-specific instrument to the psychometrically adequate instrument through the use of a statistical procedure such as correlation.
3. Identify which change scores will yield the most useful information. For example, comparisons between fall entry scores and spring exit scores may be better suited in some cases, while in others quarterly change comparisons may be preferable.
4. Use individual subject information in final analyses only if it will provide more or different information than is yielded by total or aggregate scores.
5. When dealing with small samples of children ($N<6$), individual analysis may be more profitable than aggregate analysis.

Consideration of the above points should encourage personnel to continue to focus on the documentation of child progress related to major program objectives. There should always be a clear and apparent relationship between the type of data collected, the analyses selected and program goals.

Suitability of Existing Measures

A large number of assessment measures for nonhandicapped preschool children are available. Many of these measures have been reviewed (Cross & Johnston, 1977; Walls, Werner, Bacon & Zane, 1977; Hoepfner, Stern & Nummedal, 1971). In addition, Johnson and Kopp (undated) have compiled a comprehensive bibliography of screening and measurement instruments for nonhandicapped infants. While many of the instruments described in these compendiums were designed for identification and diagnosis, early interventionists often use them to document child progress. Unfortunately, the further child performance diverges from

---

[2] A psychometrically sound instrument is one with reported adequate reliability and validity information that has been assembled in an objective manner.

normal development, the less applicable for documentation of progress are such measures. Use of these measures is often inappropriate because documentation of progress requires strategies which are comprehensive yet responsive to relatively small changes in child behavior, a characteristic not found in many of the assessment instruments that were developed primarily for nonhandicapped children (e.g., Bayley Scales of Infant Development).

In addition, many available measures have little data reflecting a handicapped population's performance on the instrument. Having data only on nonhandicapped children is appropriate if an instrument is to be used to identify children with possible problems, or to diagnose the existence of a problem. However, to chart meaningfully the progress of handicapped or at-risk children, some concept of the expected rate of progress for those children is essential.

The current lack of suitable evaluation instruments has led many interventionists to develop "home-made" tests by taking items from existing measures. This approach seems to be somewhat shortsighted. Such idiosyncratic instrument development seems likely to compound rather than to solve existing problems because most home-made instruments are not used widely enough to collect adequate information on their psychometric properties (e.g., their validity and reliability).

Until more appropriate evaluation measures become available for use with handicapped populations, we suggest that the following alternatives be considered:

1. Have a clear idea of the expected developmental goals of enrolled children. Recognize when the goal is normal development in all areas (e.g., at-risk infants), when the goal is normal development in some areas and minimal development in other areas (e.g., motor impaired child with normal functioning in social, communication and cognitive domains), or when the goal may be extremely limited change in the behavioral repertoire across domains (as in the case of profoundly handicapped children). Establishing individual and group goals should guide the search for selecting the most appropriate instruments (Sheehan, 1979; Harbin, 1979).
2. When possible, select instruments which have known reliablity, validity and normative data.
3. Use a battery of instruments which reflect desired program goals. If possible, include standardized instruments in the battery to relate them to non-standard instruments. For example, examine the relationship between a child's performance on UPAS and on the McCarthy Scales of Children's Abilities (McCarthy, 1972).
4. When no standard instruments are available that match program goals, define in clear behavioral (objective) terms what those goals are, then use those definitions as the basis for developing an assessment measure.

The above suggestions are compromise solutions but must suffice until the field develops more appropriate instruments for measuring the progress of young handicapped infants and children.

Appropriateness of Existing Analytic Designs

Use of traditional research designs confronts the early interventionist with serious problems. Randomly assigned control groups are generally not possible (Bricker, 1978). Quasi-experimental designs offered as substitutes for such control groups still generally require comparison subjects who receive no treatment and therefore are often equally impossible (Cook & Campbell, 1979; Bricker & Sheehan, 1980). Nevertheless, two alternatives are available to contend with the problem of no control or comparison groups. The first

alternative involves the use of analytic designs, which depend on statistical comparison groups. The second alternative involves comparisons using other existing intervention programs.

The general model using statistical comparison groups has been labeled either a prediction model (Sheehan, 1980) or a norm-referenced model (Horst, Tallmadge & Wood, 1975; Tallmadge, 1977). In essence, this analytic model involves the computation of predicted progress for handicapped children and the comparison of actual progress with that prediction. Predicted progress can be derived in several ways such as reference to existing norms, correlations between age and pretest, and correlations between posttest of one year and posttest of the next year.

Each of these analytic models faces problems stemming from assumptions of linear growth and stable estimates of progress (Linn, 1979). Growth of handicapped children (or even nonhandicapped children) does not necessarily conform to those assumptions. The effects of deviations from assumptions of linear growth and stable progress on the analytic designs is unclear. Nonetheless, intervention programs should consider the incorporation of prediction designs, rather than rely exclusively on a pretest/posttest comparison with no estimate of progress in the absence of intervention.

A second alternative for non-comparison controls is to compare child progress in a given intervention program to child progress in other intervention programs. If interventionists are willing to compare child progress across programs, such designs are possible. This possibility will become more appealing as the question "Is intervention better than no intervention?" is replaced with "Which intervention is most effective?"

In the development of an evaluation plan, we suggest the following strategies be used to deal with problems posed by design issues:

1. If possible, use an instrument for which normative data exist. Using an instrument with norms, one can derive expectations of progress on other, non-normed, instruments through correlation of the two instruments. For example, a comparison can be made of children's performance on the Bayley with their performance on a non-standardized checklist. If a reasonable correlation is found between the two instruments, then one can place more reliance on the checklist as an accurate reflection of child progress. Keep in mind, however, that a low correlation does not necessarily mean the checklist is inappropriate.
2. Be prepared to modify expectations of progress if the derived expectations consistently underestimate or overestimate actual child progress. For example, historically the predictions of progress for Down's syndrome children were bleak; now outcomes for these children are more optimistic, and expectations must be modified accordingly (Hanson, 1977).
3. Use complex analyses sparingly. Other evaluation efforts have shown that statistical complexity can result in limited generalization and interpretation (Haney, 1977).
4. If possible, make arrangements with other intervention programs to share data bases. Be sure to establish procedures for guaranteeing anonymity of children, parents and projects.
5. Consistent patterns of performance across tests, conditions and children lend validity to the general nature of the results. Therefore, evaluate child progress using multiple indices and conditions when possible; for example, include standardized tests, criterion-referenced tests and data collected in different settings.

Documenting Child Progress: An Evaluation Plan

The design of a plan documenting child progress necessitates the adoption of a pragmatic perspective. The remainder of this paper describes the documentation of child progress in the Preschool Program at the Center on Human Development, University of Oregon. The evaluation strategies were selected as a balance between methodological purity and the demands of field research in a service delivery system. The detailed description of this evaluation plan for child progress may serve as a useful model, in whole or in part, for other early intervention programs. The evaluation plan described below is composed of four major sections that are consistent with the problem areas previously discussed. The four sections are: resources, framework, measures and analyses.

Resources

Location of this program in the Center on Human Development has fostered the combination of a variety of resources to benefit current research, training and demonstration efforts. For example, providing graduate training in the area of early childhood special education necessitates placement of students in sites involved in the active delivery of educational and allied health services to handicapped infants and children. Preferably, the staff at these sites are engaged in effective intervention approaches as well as in investigating innovative strategies which may ultimately improve intervention efforts. While the demonstration program provides students with practical experience, the students in turn provide the program with the valuable resource of an additional workforce to direct individual and group activities for the enrolled children.

The research activities have been defined to include documentation of child progress as an important goal; therefore, rather than operating the research project independently of the demonstration unit, the demonstration unit has become one of the focal points for the research staff. In particular, the research staff has assisted in preparing the outcome data for analysis and in conducting subsequent analysis, largely using prepackaged computer programs.

The state agency, Crippled Children's Division, which is primarily responsible for evaluating and diagnosing developmentally disabled children in Oregon is also housed in the Center. This proximity has enhanced a cooperative liaison in which the CCD completes the initial evaluation and diagnostic activities on a child, leaving the intervention staff free to focus their evaluation efforts on educational assessments.

Framework

The theoretical framework we have developed for evaluation is based on three tenets: 1) behavior changes in form from the simple to the more complex; 2) disequilibrium produced by changing environmental demands is necessary for eliciting new adaptive responses; and 3) behavior develops sequentially following general but consistent guidelines.[3]

---

[3]We are well aware that handicapped infants and children often show behavioral deviations produced by a specific disability (see Fraiberg, 1968, or Kopp & Shaperman, 1973). These

This developmental-interactive approach, a combination of developmental theory and behavioral technology, provides a general map of emerging behavior. These maps are based on data generated by the developmental literature in the domains of motor, sensorimotor, communication and social behavior. Such developmental hierarchies are composed of a series of sequentially acquired behaviors and currently provide our best criteria for establishing objectives for child programming in all but the most severely handicapped individual. This framework not only allows us to specify long-term objectives but assists us in determining the most probable sequence for successful programming.

This developmental-interactive framework with its attention to developmental requisites or concurrent responses can suggest what the immediate intervention priorities should be. Sensitivity to relationships among domains of behavior can help the interventionist select targets that are appropriate both within and across domains. For example, many early social and self-help skills require a level of understanding of objects in terms of their social utility (e.g., spoons are for eating, shoes are for wearing). To understand these functions, it appears that the child must have passed the stage where objects are only sucked, banged or dropped. The child must be alert to the unique physical properties of objects before he or she can begin to understand their social significance.

The developmental-interactive framework indicates strategies of intervention that will foster development toward the chosen program objectives. Beyond suggesting a task analysis, the information available suggests strategies which should, if applied correctly, produce more functional and universal response forms. For example, the interventionist operating within this framework will not focus exclusively on verbal expression when working on communication but will also target important sensorimotor and conceptual skills that research suggests underlie or coincide with early language acquisition. This framework can also provide a basis for selecting evaluation instruments that specify reasonable developmental sequences. Such sequences should lay out objectives according to the processes that probably underlie the target areas and suggest the interrelationships among target domains as well.

Finally, the developmental-interactive framework can direct the modifications and adaptations of curricular materials for individual children as well as guide the construction of new materials. Developmental sequences taken from the literature can only provide a general set of bench marks. Within this general framework, the interventionist must be prepared to make numerous modifications to meet the needs of young disabled children who may vary significantly from the developmental patterns typically followed by nonhandicapped children. That is, a general sequence of development may suggest that infants learn to roll over, crawl, pull to stand, and then take independent steps. A motorically handicapped child may not be able to crawl or pull to stand, but the goal nonetheless is independent mobility for that child if at all possible, even though the interventionist may have to substitute some form of support for the child to be able to stand and walk for moving on a crawling board.

---

deviations appear primarily to be tied to the surface responses rather than to the conceptual or superordinate level. Hearing-impaired children may not learn oral communication, but they do acquire the ability to send and receive communicative messages. So too motor-impaired children may never walk without support, but most likely they will acquire some means for ambulating. General maps of development can provide useful guidelines for most children except, perhaps, for the most severely disabled.

Measures

Assessment was directed toward collection of information and monitoring of child progress in four distinct but related areas: demographic, norm-referenced, criterion-referenced and curricular. Table 1 contains the areas of evaluation included, the specific instrument, frequency of administration and expected outcomes. Table 2 contains the testing schedule used in the program. Appendix A contains brief descriptions of the assessment instruments shown in Tables 1 and 2. The paragraphs which follow describe the assessment methods in each of the four areas.

Demographic Assessment. Upon entry into the program, a demographic form (Appendix B) was completed on each child and his or her family. This brief form contains items which have been constructed to yield quantifiable responses. Areas covered included: identification and description of the child and family, prenatal data, information regarding child's impairments and handicapping conditions, and findings of previous assessments.

Demographic forms were updated each fall at the beginning of the school year or when any children entered the program. The information collected from the demographic data forms (minus identifying information) was then transferred to FORTRAN sheets, along with code numbers. The data were then key punched onto computer cards to be used in subsequent analyses.

Norm-Referenced Assessment. As indicated earlier, no ideal or completely appropriate instruments are available for use with populations of handicapped infants and preschoolers. In recognition of this deficit, interventionists must make compromises in selecting norm-referenced instruments. This project, therefore, chose two instruments which have norms available for nonhandicapped children. These include the Bayley Scales of Infant Development (Bayley, 1969) and the McCarthy Scales of Children's Abilities (McCarthy, 1972) (see Appendix A).

The Bayley Scales and the McCarthy Scales were administered during September and October or upon entry, and again in May through June or upon exit. The staff psychologist administered these tests according to the developmental age schedule contained in Table 2.

It should be noted that an overlap exists in test administration of the Bayley Scales and the McCarthy Scales. With nonhandicapped children, the Bayley Scales are considered to be suitable for children up to 30 months of age, and the McCarthy Scales are considered to be suitable for children exceeding 30 months of age. These considerations must be altered for handicapped children, since the performance of many handicapped children can be scored on the Bayley Scales after 30 months of age, when performance cannot be scored on the McCarthy Scales. The staff of the intervention program decided to administer the Bayley Scales until a child functioned at a developmental age of 30 months. At this point, the McCarthy Scales were also administered until a child achieved a scorable General Cognitive Index (GCI) on the McCarthy Scales, at which point the Bayley Scales were no longer used. This procedure resulted in a smooth transition in documenting handicapped children's progress between the developmental periods measured by the Bayley Scales and the McCarthy Scales.

The norm-referenced data were coded on FORTRAN sheets according to a specified format. Computer entry was directly derived from test protocols to avoid error.

Criterion-Referenced Assessment. Three criterion-referenced instruments were used in the intervention program. These included the Student Progress Record (SPR), the Uniform Performance Assessment System (UPAS) and the Adaptive Performance Instrument (API) (see Appendix A). Differences between norm-referenced assessments and criterion-referenced instruments are not always clear (Tallmadge & Horst, 1976); however, there are distinctions. Criterion-referenced instruments generally have more items than norm-referenced instru-

Table 1

Description of Instruments Used in the Evaluation Plan

Areas of Evaluation	Instrument	Frequency of Administration	Outcomes
mographic Information	Preschool Demographic Form	Entry into program; Beginning of year; Exit from the program	Quantifiable information on child and family
rm-Referenced Data	Bayley Scales of Infant Development (Bayley, 1969) McCarthy Scales of Children's Abilities (McCarthy, 1972)	Entry into program; Beginning and end of school year; Exit from the program	Raw scores, Scale scores Developmental Quotient, or Age Equivalent
iterion-Referenced ta	Student Progress Record (Oregon State Mental Health Division, 1977) Uniform Performance Assessment System (White et al., 1978) Adaptive Performance Instrument (CAPE, 1978)	Beginning and end of school year Quarterly Quarterly	Percent of items passed
rriculum-Linked ta	Preschool curriculum covering the following areas: Communication Fine and Gross Motor Sensorimotor or Preacademic Social/Self-help	Weekly probes	Individual progress data toward objectives

Table 2

Testing Schedule

Instrument	Developmental Age			
	Birth-2 years	2-3 years	3-4 years	4-5 years
Bayley	^----------^			
McCarthy		^----------^	^----------^	^----------^
Adaptive Performance Instrument	^----------^			
Uniform Performance Assessment System		^----------^	^----------^	^----------^
Student Progress Record	^----------^	^----------^	^----------^	^----------^
Curriculum Monitoring	^----------^	^----------^	^----------^	^----------^

ments which tend to make criterion-referenced approaches comprehensive rather than representative. Further, the content of criterion-referenced instruments often more closely approximates the objectives of intervention than does the content of the norm-referenced measures. Items in criterion-referenced, developmental instruments are generally chosen because of their apparent value as educational training targets. Items in the norm-referenced instruments are chosen for their predictive value, their presumed importance in normal development and their ease of administration. Norm-referenced items are not usually chosen for specific educational relevance.

Criterion-referenced measures provide norms or points of comparison, just as the norm-referenced measures do. For example, two of the three instruments noted, UPAS and SPR, provide comparative data; the third instrument, API, is currently being field tested, from which some initial standardization data will be forthcoming. Establishing norms is not the major advantage of the criterion-referenced measures; the strength of these instruments lies in the relationship which their content has to specific program goals and objectives.

The Student Progress Record (SPR) is a developmentally based instrument that covers 14 important areas of behavior (Oregon State Mental Health Division, 1977). The teacher or home interventionist administered this test in the fall and again in the spring. The program's instructional staff administered the Uniform Performance Assessment System (White, Edgar & Haring, 1978) and the Adaptive Performance Instrument (CAPE, 1978) quarterly according to the developmental age schedule specified in Table 2.

Curricular Assessment. In the fall, an IEP was written for each child in which parents and staff specified long-term and quarterly short-term goals to be used in curricular assessment. To insure a comprehensive intervention program for each child and to insure continuity for a child from year to year, a set of programmatic training targets were developed for the following domains: language, speech-hearing, gross motor, self-help, pre-academic, fine motor and sensorimotor. Each domain was divided into a number of sequential targets that were appropriate for the majority of children enrolled in the program. The targets were laid out developmentally, and as often as possible, the intervals between targets approximated two to three months. As an example, Table 3 contains a list of the long-term training targets, birth through 36 months, for the fine motor domain.

If an infant were developing normally, one would expect that all five fine motor targets specified from birth to 12 months in Table 3 would be accomplished during the first year, while mildly to moderately handicapped infants would be expected to acquire fewer targets, and the more severely impaired infant, still fewer targets per year. The same procedure was used to determine long-term targets across domains. Following each school year, a new IEP was written for each child incorporating a new set of long-term training targets. If the infant, for example, had completed the first three targets in the fine motor area, items 4 through 7 might be included as targets for the coming year. Such a system provided a general continuity for the selection of IEP objectives from year to year for each child. After targets were selected and arranged by priority, a planning sheet (Table 4) for the chosen targets was completed.

To insure consistency and continuity in monitoring child progress, systematic, data-collection activities were necessary. To assist the instructional staff, three types of data-recording forms were devised. These forms were designed to collect trial-by-trial data, multiple targets or subjects data and continuous data.

Table 3

Training Targets for the Fine Motor Domain

Age in Months	Training Targets	
	Skill	Activity
0-12	Eye tracking visually directed	1. Tracks 180 degrees, both directions.
		2. Grasps objects held at all levels.
	Voluntary release	3. Releases block into cup.
	Midline orientation	4. Manipulates objects with both arms about midline.
	Transfer of object	5. Uses both hands to transfer objects.
12-24	Pincer grasp	6. Picks up items with pincer.
	Wrist rotation	7. Turns lid on jar; opens door.
	Refined pincer grasp	8. Turns pages in book.
	Eye-hand coordination	9. Builds tower of 2-3 blocks.
	Visual Motor	10. Places 3 shapes in form board.
24-36	Eye-hand coordination and pincer grasp	11. Places 2 small pegs in peg board.
	Eye-hand coordination and pincer grasp	12. Imitates vertical and horizontal lines.
	Eye-hand coordination and pincer grasp	13. Imitates circular lines.
	Visual Motor	14. Uses scissors to snip paper.
	Visual Motor	15. Completes 3-piece puzzle.

Table 4

A Sample Planning Sheet

PLAN SHEET

Plan Sheet number _____ of _____ Date _____ Manager (person implementing program) _____

Pupil ___(child's name)_____ STOBJ (relevant short-term objective) _____ LTOBJ (relevant long-term objective) _____

Target (area working in) _____ STO completion date (projected) _____ LTO completion date (projected) _____

	Acceleration		Deceleration	
Antecedent	Response	Consequence	Response	Consequence
(This includes events set up to bring about objective and what the teacher needs to do and set up before she can ask for the targeted behavior.)	(This is a description of the targeted behavior--behaviorally written with criteria specified.)	(This is a description of events set up to consequate the targeted behavior.)	(This is a description of behaviors that are incompatible with the targeted behavior and that may interfere with the teacher's implementation of the program.)	(This is a description of events that are set up to move each incompatible response closer to the targeted behavior.)

Table 5

Sample Recording Form for Trial by Trial Data

Target Area: Fine Motor

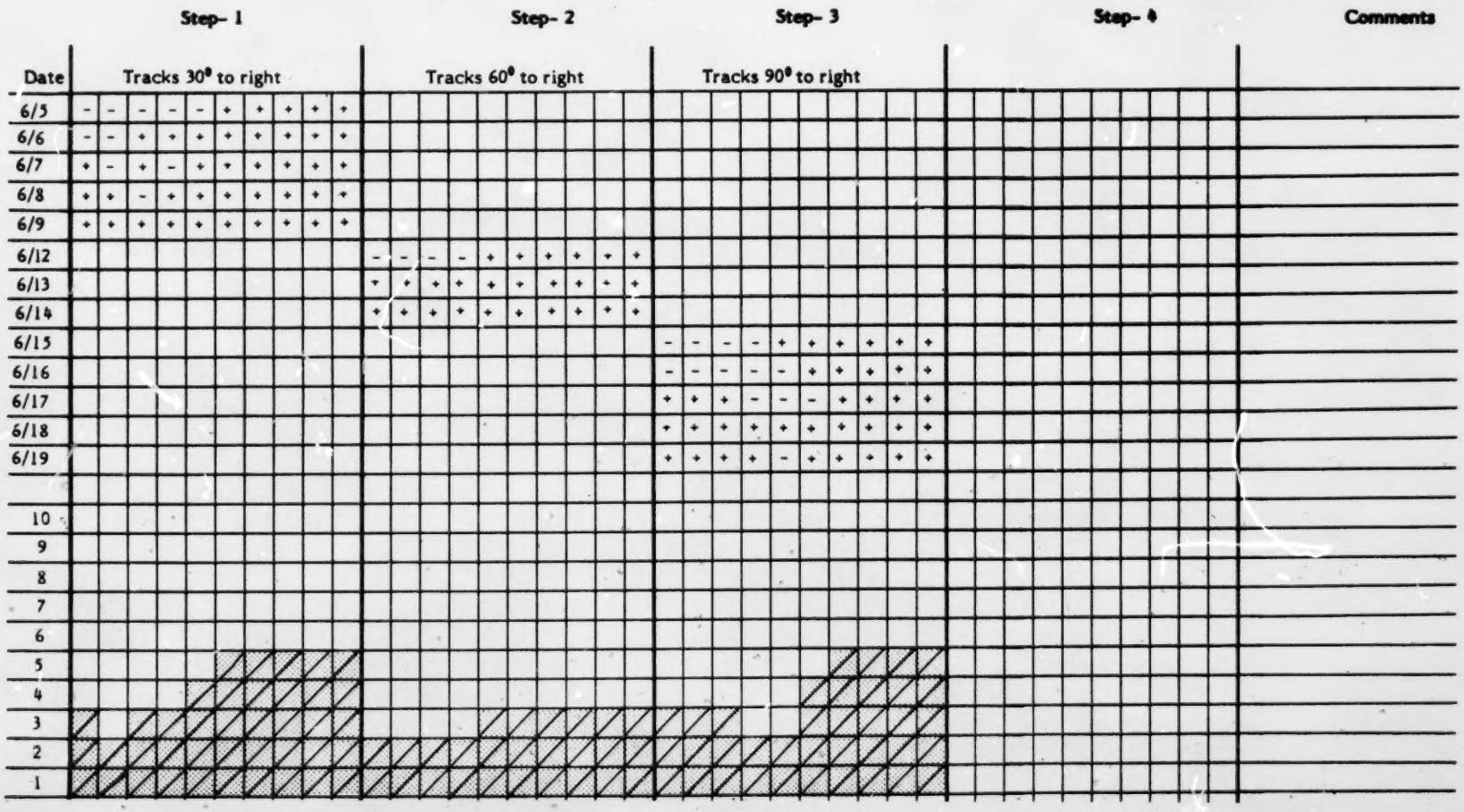
LTO: Infant will track a slowly moving object 180° in both directions.

Original STO: Infant will track from mid-line to right.

Data Collection Specifications: Ten trials per day; criteria 8/10 trials correct for two consecutive days

Program Objective: Infant will track 30°, 60°, and 90° from mid-line to right.

Child: A Teacher: Jones



The data-recording form contained in Table 5 was designed to collect trial-by-trial data and also to provide space for indicating the following information:

Target Area	the general area within which the objectives are located, such as gross motor
LTO	the long-term objective
Original STO	the original short-term objective from the Programming Training Targets (in the event that a teacher must branch from the STO in the skill sequence, this space specifies that STO from the sequence. In this way, the data records always relate to the core skill sequences, in spite of necessary variations by each teacher.)
Program Objectives	either the STO from the skills sequence (if it was appropriate), or the objective that was branched from the original STO
Data Collection	the specific data-collection procedures: e.g., 15-minute time sample, frequency data

This data-recording form is composed of three sections: a space to indicate the data, a series of columns to record trial-by-trial data, and a space for comments. Data columns are divided into four sections, each composed of ten squares. The space above the recording squares is for specifications of the behavior, cues and criterion for that particular target step in the training program. Moving from left to right, steps targeted should more closely approximate the program objective. This form can be used to record a variety of data, such as frequency, correct/incorrect, rate or interval. The bottom of the form can be used to graph the data.

Table 6 contains a form for recording data on multiple subjects or multiple targets for one child. This form is particularly useful for collecting simultaneous information on small groups of children or for comparing an individual child's progress across four different targets. If a teacher wants to increase the number of trials per session, the form is easily modified.

The recording form contained in Table 7 was developed to accommodate the collection of continuous data in which it is important to specify the context, antecedent and consequences for a targeted behavior, as when monitoring the occurrence of productive language during a free play period.

In general, data were collected on a child's progress toward targets on a daily or weekly basis. The frequency of data collection was dependent upon a number of factors.

Analyses

The majority of analyses conducted on the data generated from this program involved the use of the University of Oregon's computer facilities. Use of computers for data analysis requires that care be taken to protect the confidentiality of children and families served by the intervention project. To insure privacy, code numbers were assigned to children and families, and these code numbers were the only identifying information computerized. A roster of code numbers with identifying information was kept in a secure location. This roster had only restricted access.

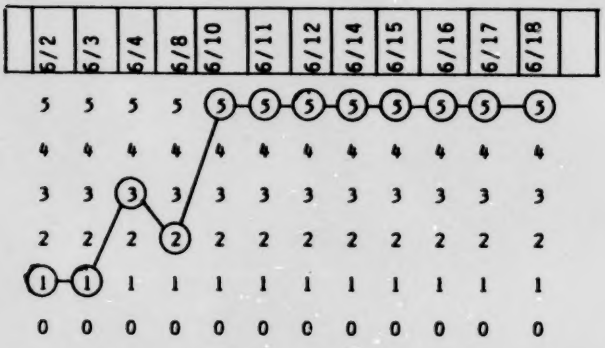
Table 6

Sample Recording Form for Multiple Subjects or Targets

Name: A

Target: Answering questions appropriately during small group time

Steps: [1]Sitting in chair; [2]Attending to teacher; [3] responds appropriately when asked a question.

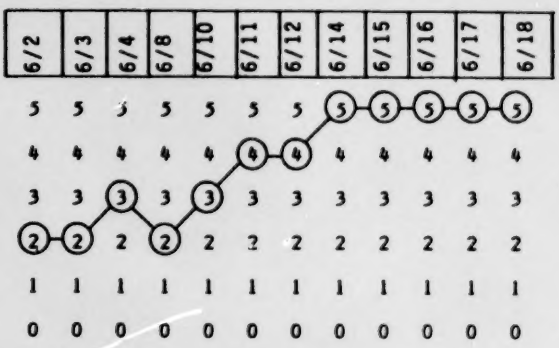


Baselines:

Name: B

Target: same

Steps:

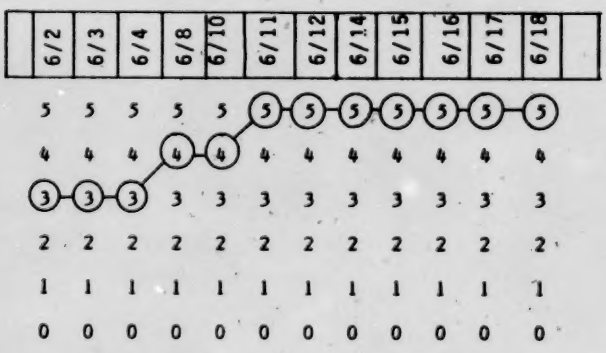


Baselines:

Name: C

Target: same

Steps:

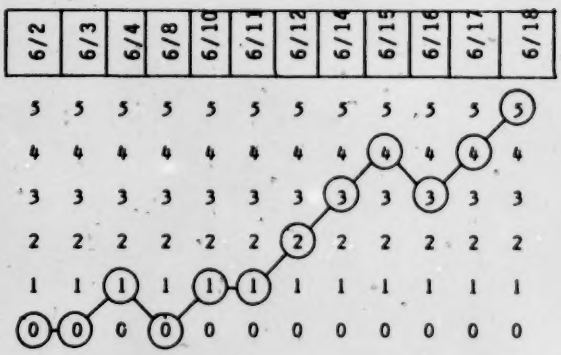


Baselines:

Name: D

Target: same

Steps:



Baselines:

* Circled number indicates number of correct responses.

Table 7

Sample Recording Form for Continuous Data

Domain: Communication

Child: A

LTO: Expresses linguistic functions of labeling of objects

Teacher: Dan

Program Objective: responds when asked to label object

DATE	CONTEXT	ANTECEDENT	BEHAVIOR							CONSEQUENCE
			Vocalize	Gesture	Word					
1/20	Play area/peer	Peer points			ball					Peer takes ball
1/20	Play area/peer	Ball rolls away			go					Peer looks at ball
1/20	Play area/peer	Ball rolls away		points	ball go					Peer looks at ball
1/20	Play area/peer	Ball rolls away		points	ball go					Peer retrieves ball
1/20	Play area/peer	Peer holds ball		points	ball					Peer gives ball to Kevin

Data Preparation

One problem that plagues any complex data-gathering effort involves errors occurring at the time of assessment, during transfer to coding sheets, during key punching and during computation. Prior to the formal data analyses, the data collected by the project personnel were checked for possible error. Because this project did not have the necessary resources to double-check all data collection and processing, the following strategy was adopted to insure that data were free from errors which might distort the analyses. After the data were keypunched and verified by the keypunch operator, each variable was processed through the FREQUENCIES PROGRAM of the Statistical Package for the Social Sciences (Nie, Hull, Jenkins, Steinbrenner & Bend, 1975). The frequency analysis produced by the SPSS program was visually examined to determine whether the range, mean and general distribution patterns were reasonable, given the nature of each variable. Variables with possible errors were identified and rechecked.

Following these preliminary checks, a computerized list of all cases (subjects) for each variable was generated using the LIST CASES PROGRAM of SPSS. The list case tables were used to identify cases with questionable data. These data were checked and errors corrected as necessary. Such a procedure may need to be repeated in order to identify all possible errors.

Simple Data Description

The next analysis undertaken focused on generating descriptions of the data. Using the LIST CASES routine of SPSS, tables of the raw data for each child and aggregated summaries of means, standard deviations, ranges, etc., were generated for the total group and subgroups of normal, at-risk, mildly handicapped, moderately handicapped and severely handicapped children. Such tables can be compiled by hand if computer facilities are not available.

This descriptive analysis was conducted to present a case-by-case description of all the relevant data with accompanying group statistics to obtain a visual estimate of the extent to which the group data were consistent with the individual data. When children's individual data are not similar, or at least not evenly distributed, those data should not be aggregated. It may be apparent from a descriptive analysis, for example, that approximately one-half of the children are regressing, while the other half are progressing. If this is so, each subgroup should be analyzed separately. Otherwise, a simple averaging process will indicate no change, clearly an inaccurate indication of the actual performance of children in the program. Although this process is time-consuming, it is essential that case-by-case visual examination of the data be conducted to determine whether subsequent, subgroup analyses are needed and appropriate. Our analysis revealed no such discrepancies, so data were analyzed for total groups and relevant subgroups.

The simple data-description step in the analytic process is also useful for another purpose. At this juncture, the group statistics of mean, range, median and standard deviation can provide an initial look at average performances at each time period. (These statistics can be generated by the SPSS program, CONDESCRIPTIVE.) Due to the possible need for subgroup analyses, the average performances might be somewhat distorted, but where such analyses are appropriate, they can be examined and reported. (Any distortions in distributions, such as skewness, are also indicated by the SPSS program CONDESCRIPTIVE.)

This simple data-description procedure provided a perspective which was useful in guiding the staff in conducting further analyses. Decisions were made to reduce the quantity of data used in subsequent analyses. For example, no further analyses were conducted on the subtest scores of the McCarthy. Such decisions should be made when a review of the analysis

reveals that the subtest scores exhibit little variation from each other or fail to contribute additional information. Subtests or total scores can also be identified which exhibit so little variation that it is unproductive to perform further analysis. The presence of strong ceiling or basal effects is another reason to forego further analysis. Ceiling and basal effects occur when scores cluster at the high or low end of a scale. Such clustering obscures possible differences between scores which would have been found if the scale's range had been greater. In our analysis of the children's performance on the Bayley Scales, for example, we found that a sizeable proportion of the subjects scored below 50. Scores of this nature prohibit computing a Mental Development Index or Psychomotor Development Index, and therefore we did not attempt these comparisons with our population of handicapped infants and toddlers.

Estimation of Program Effects

Program effects on the children enrolled in this project were determined in three ways. The first method was to compare pretest (fall) scores to posttest (spring) scores, using a correlated t-test [4] to determine whether there was a statistically significant difference.

A second strategy involved determining whether the differences between pre and posttest scores were educationally significant. Establishing educational significance can be achieved in two ways. Abt Associates (1977) and Tallmadge (1977) have suggested a procedure in which gains from pre to posttesting are compared with the average, pooled standard deviation of the pretest and posttest scores. If they exceed .25 (Abt Associates) or .33 (Tallmadge), the gains are considered to be educationally significant. In addition, educational significance can be established through a discussion of the relative meanings (as determined through the literature) of the magnitude of the obtained mean differences. For example, the average difference occurring during a 6-month interval may be 8 months gain. A review of the literature may reveal that such gain occurs only rarely in a specific subpopulation of handicapped children; thus, these 8-month gains would be educationally significant.

The third method chosen to examine program effects involved comparing gains made by this population of children on the SPR with those made by a state-wide control group. The

---

[4] It is useful to describe the roles that inferential and descriptive statistics should play in the documentation of intervention effects. Pre/posttest differences are commonly assessed with the correlated t-test. This test is designed to determine whether differences found between related samples drawn from a large population indicate differences in their respective populations. Unfortunately, intervention programs do not actually sample from clearly defined populations; therefore, it is inappropriate to use the correlated t-test as a method of inferring differences in larger populations. Inferential statistics, however, are often used to document differences due to intervention. The hidden assumption is that, if appropriate sampling had been done, the results would have generalized to the larger population. In a sense, the results of inferential statistical procedures such as the t-test may be thought of as descriptive rather than inferential. The difference is about the target group of children, not about some larger population from which a random sample was drawn. This points up the value of using other descriptive statistics, such as means, modes, standard deviations and the like, as ways of describing the target group and how the group changed as a result of the intervention. A pre/posttest correlated t-test may be a useful descriptive tool for documenting differences, but since it is not generally used as an inferential statistic in intervention research, results should be generalized cautiously.

gain scores of the project's three-, four- and five-year-old- children were compared with gains made over the same period of time by other three-, four- and five-year-old children attending other intervention programs throughout Oregon. A number of similar strategies are available to interventionists. Descriptions of these procedures are contained in Appendix C.

Figure 1 presents a synopsis of the information presented on instrument selection, data preparation and analyses. This figure presents an outline of the main considerations involved in moving from instrument selection to data analysis.

Summary

The Preschool Program, Center on Human Development, has evolved a comprehensive and organized evaluation plan which was initiated with the careful selection of a battery of instruments that included standardized and criterion-referenced measures. These measures were selected because program resources were available for implementation and, more importantly, because they reflected program emphasis. The measures were administered according to a preconceived plan, and the data were checked and properly stored until the analyses were completed. Prior to analyzing the results, the data were inspected for error. A straightforward statistical analysis was adopted to compare the pretest with posttest performance across all measures. In addition, educational significance was measured. The outcome of this evaluation plan has been presented in detail elsewhere (Bricker, 1981; Bricker & Sheehan, 1980).

Documentation of child progress is a complex task, and the field of early childhood/special education is still struggling to develop appropriate and usable assessments and methodological strategies to determine program impact. Recognizing this deficit, this paper has been formulated to present an evaluation plan that may enhance the documentation efforts of the field.

Figure 1

Flowcharts of Important Considerations in Instrument Selection, Data Preparation and Data Analysis

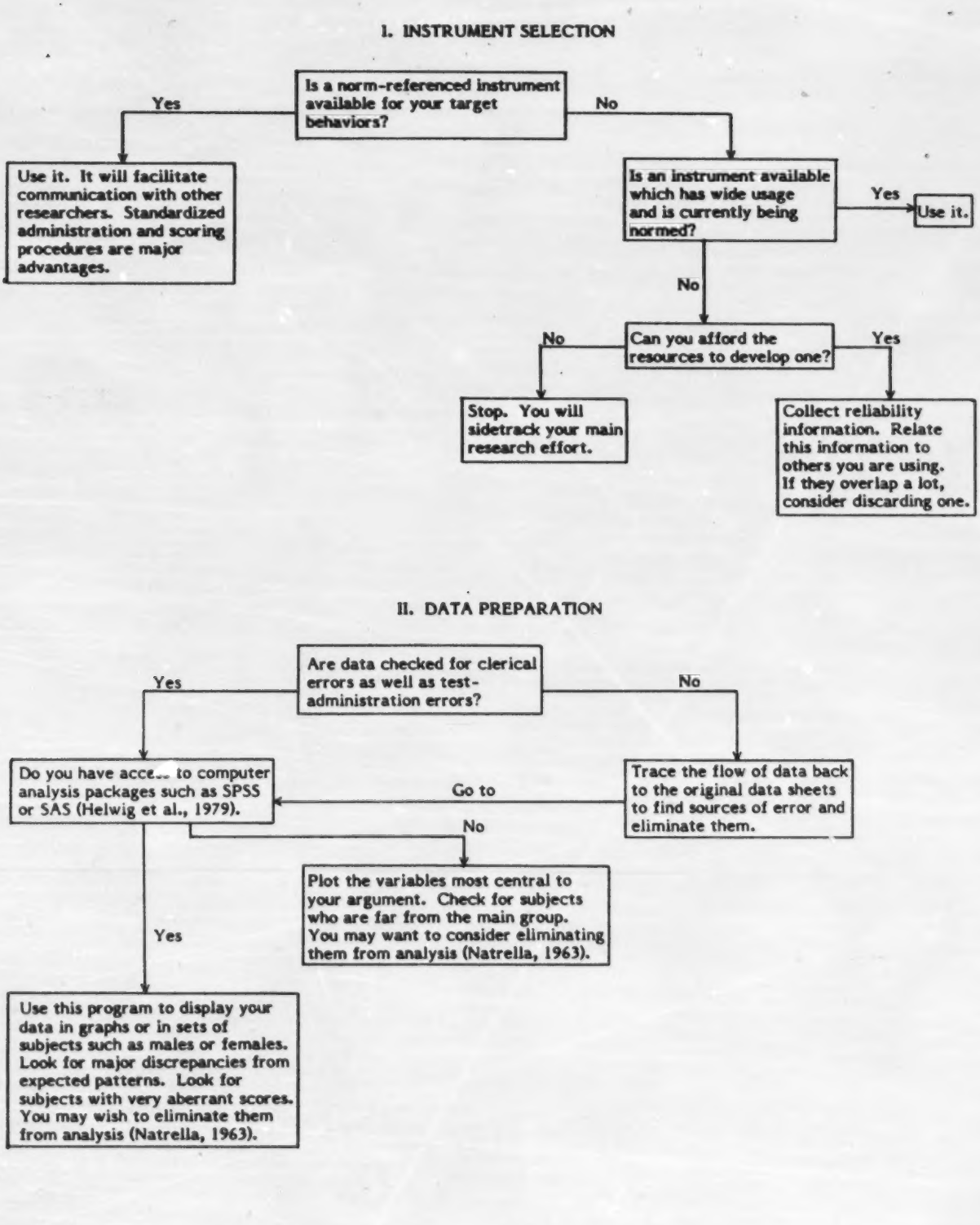
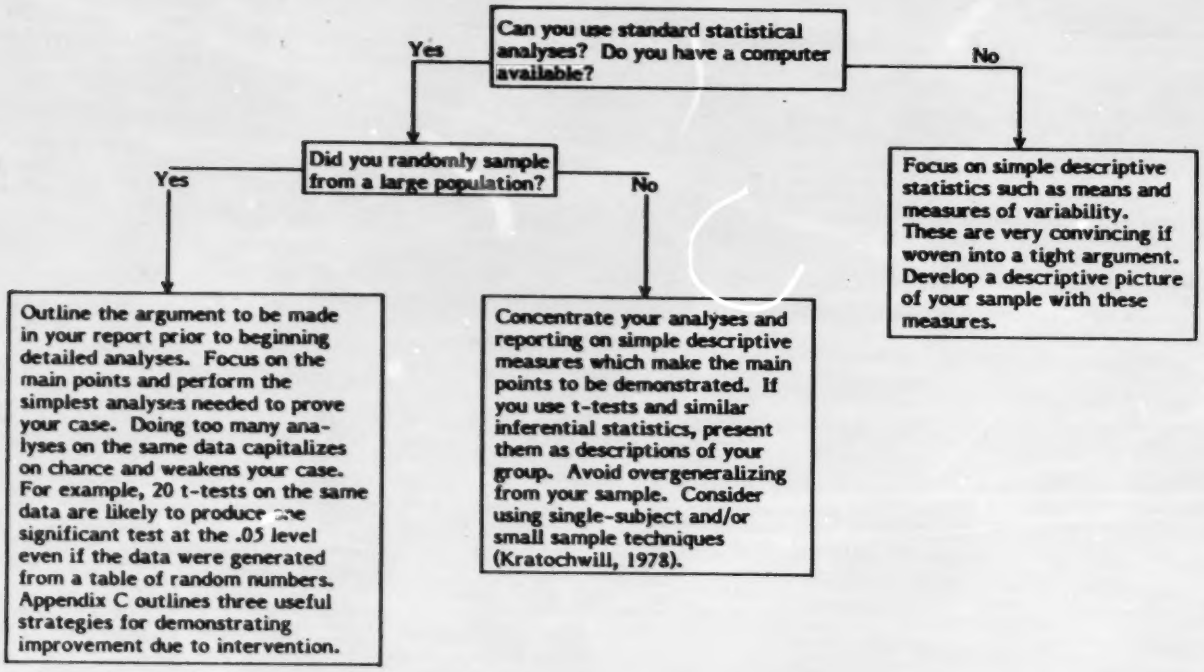

I. INSTRUMENT SELECTION
Is a norm-referenced instrument available for your target behaviors?
Yes
No
Use it. It will facilitate communication with other researchers. Standardized administration and scoring procedures are major advantages.
Is an instrument available which has wide usage and is currently being normed?
Yes
Use it.
No
Can you afford the resources to develop one?
No
Yes
Stop. You will sidetrack your main research effort.
Collect reliability information. Relate this information to others you are using. If they overlap a lot, consider discarding one.
II. DATA PREPARATION
Are data checked for clerical errors as well as test-administration errors?
Yes
No
Do you have acce... to computer analysis packages such as SPSS or SAS (Helwig et al., 1979).
Go to
Trace the flow of data back to the original data sheets to find sources of error and eliminate them.
No
Plot the variables most central to your argument. Check for subjects who are far from the main group. You may want to consider eliminating them from analysis (Natrella, 1963).
Yes
Use this program to display your data in graphs or in sets of subjects such as males or females. Look for major discrepancies from expected patterns. Look for subjects with very aberrant scores. You may wish to eliminate them from analysis (Natrella, 1963).

Figure 1 (continued)

III. DATA ANALYSIS


Can you use standard statistical analyses? Do you have a computer available?
Yes
No
Did you randomly sample from a large population?
Yes
No
Focus on simple descriptive statistics such as means and measures of variability. These are very convincing if woven into a tight argument. Develop a descriptive picture of your sample with these measures.
Outline the argument to be made in your report prior to beginning detailed analyses. Focus on the main points and perform the simplest analyses needed to prove your case. Doing too many analyses on the same data capitalizes on chance and weakens your case. For example, 20 t-tests on the same data are likely to produce one significant test at the .05 level even if the data were generated from a table of random numbers. Appendix C outlines three useful strategies for demonstrating improvement due to intervention.
Concentrate your analyses and reporting on simple descriptive measures which make the main points to be demonstrated. If you use t-tests and similar inferential statistics, present them as descriptions of your group. Avoid overgeneralizing from your sample. Consider using single-subject and/or small sample techniques (Kratochwill, 1978).

Appendix A

Description of Assessment Instruments

Bayley Scales of Infant Development (Bayley, 1969)

The Bayley Scale is a tripartite developmental instrument designed to be administered during the first 2½ years of life. The three parts are considered to be complementary and include:

(1) The Mental Scale is designed to assess sensory-perceptual acuities, discriminations and the ability to respond differentially to different objects and events; the early acquisition of "object constancy" and memory, learning and problem-solving ability; vocalizations and the beginnings of verbal communication; and early evidence of the ability to form generalizations and classifications, which is the basis of abstract thinking. Results of the administration of the Mental Scale are expressed as a standard score, the MDI, or Mental Development Index.

(2) The Motor Scale is designed to measure the degree of control of the body, coordination of the large muscles and finer manipulatory skills of the hands and fingers. As the Motor Scale is specifically directed toward behaviors reflecting motor coordination and skills, it is not concerned with functions that are commonly thought of as "mental" or cognitive in nature. Results of the administration of the Motor Scale are expressed as a standard score, the PDI, or Psychomotor Development Index.

(3) The Infant Behavior Record (IBR) is completed after the Mental and Motor Scales have been administered. The IBR helps the clinician assess the nature of the child's orientations toward his or her social and material environment as expressed in attitudes, interests, emotions, energy, activity and tendencies to approach or withdraw from stimulation.

McCarthy Scales of Children's Abilities (McCarthy, 1972)

The McCarthy Scales were designed as a follow-up measure to the Bayley Scales of Infant Development. These scales are appropriate for children from 2½ to 8½ years of age. The test consists of 18 subtests that make up six scales: Verbal, Perceptual-Performance, Quantitative, Memory, Motor and General Cognitive. The General Cognitive scale is a composite of the Verbal, Peceptual-Performance and Quantitative scales. The General Cognitive scale has the same numerical characteristics as the "Deviation IQ."

Student Progress Record (Oregon State Mental Health Division, 1977)

The Student Progress Record (SPR) is a developmentally based instrument that covers 14 important areas of behavior (e.g., communication, self-help). The SPR was developed in Oregon and adopted by the Mental Retardation/Developmental Disabilities Office as the statewide mechanism for monitoring progress of all TMR and preschool handicapped infants and children.

Uniform Performance Assessment System (White, Edgar & Haring, 1978)

The Uniform Performance Assessment System (UPAS) consists of a fine motor/pre-academic, gross motor, communication, self-help/social and behavior management scale. Each scale is composed of developmentally sequenced items accompanied by extensive administration and criteria guidelines. A major drawback of UPAS is the limited number of items covering the developmental period from birth to 24 months. Consequently, this assessment instrument is useful primarily with children who are functioning above a developmental age of 24 months.

Adaptive Performance Instrument (CAPE, 1978)

The Adaptive Performance Instrument (API) covers eight domains of behavior including: physical intactness, reflexes and reactions, gross motor, fine motor, sensorimotor, communications, social and self-help. Each of the eight domains is divided into a number of test strands that are arranged in developmental sequences. The API concentrates on the developmental age span from birth to 24 months and provides procedures for novel modifications or adaptations in the assessment of children with specific handicapping conditions. These modifications are possible for children who are visually impaired, hearing impaired, visual/hearing impaired and orthopedically impaired. This instrument is currently undergoing field testing and is thus considered to be a research tool rather than a codified, published scale.

Appendix B

Demographic Form

Center on Human Development Preschool
Demographic Data Form

Today's Date _____ Site ___ ___

Pupil's Identification # ___ ___ ___ ___ ___ ___ ___ ___ ___

Pupil's Name _____
LAST FIRST MIDDLE

Pupil's Birthdate _____ SEX 1. ___ Female 2. ___ Male
mo day year

Pupil's Ethnic/Racial Group 1. ___ Caucasian 4. ___ Black 7. ___ Unknown
2. ___ Native American 5. ___ Polynesian
3. ___ Oriental 6. ___ Mixed

Time of Delivery 1. ___ Full Term (37-42 weeks)
2. ___ Premature (before 37 weeks)
3. ___ Postmature (42 weeks or more)
4. ___ Unknown

APGAR Score at 5 minutes 1. ___ 2. ___ 3. ___ 4. ___ 5. ___ 6. ___ 7. ___ 8. ___ 9. ___
10. ___ 11. ___ Unknown ___

Mother's Education 4 5 6 7 8 9 10 11 12 13 14 M.A. Ph.D. (Mark highest
Father's Education 4 5 6 7 8 9 10 11 12 13 14 M.A. Ph.D. level completed)
Other Primary Caregiver 4 5 6 7 8 9 10 11 12 13 14 M.A. Ph.D.
Unknown

Family Income (Annual Gross) 1. ___ under $5,000 4. ___ $15,000 - $20,000
2. ___ $5,000 - $10,000 5. ___ $20,000 - $25,000
3. ___ $10,000 - $15,000 6. ___ over $25,000
7. ___ Unknown

Impairments:

Intellectual/Cognitive:

1. ___ Normal 2. ___ Mild 3. ___ Moderate 4. ___ Severe 5. ___ Profound

6. ___ Unknown

Hearing:

1. ___ Normal 2. ___ Some Impairment 3. ___ Deaf 4. ___ Unknown

Vision:

1. ___ Normal 2. ___ Some Impairment (glasses) 3. ___ Legally Blind

4. ___ Totally Blind 5. ___ Unknown

Motor:

1. ___ Normal 2. ___ Some Impairment 3. ___ Non Ambulatory 4. ___ Unknown

Behavioral Problems:

1. ___ Normal, none 2. ___ Mild 3. ___ Severe 4. ___ Unknown

Identified Conditions:

1. ___ Down's Syndrome
2. ___ Cerebral Palsy
3. ___ Myelomeningocele
4. ___ Seizure Disorder
5. ___ Hydrocephalism
6. ___ Microcephalism
7. ___ Autism
8. ___ None
9. ___ Other _____
10. ___ Unknown

Standardized Test Score (most recent):

1. ___ Bayley
2. ___ Binet
3. ___ McCarthy
4. ___ Denver
5. ___ Cattell
6. ___ Other
7. ___ Unknown

Score: _____

Date: _____

Appendix C

Additional Analytic Strategies

For Assessing Program Impact

1) A prediction of progress can be made by computing the expected posttest scores on the basis of the child's pretest performance. The predicted posttest scores are estimations of how children would have progressed in the absence of intervention. For example, it may be predicted that a subset of children will progress three months in developmental age over a 9-month period. The actual progress of the children is then compared to the predicted progress in order to estimate a program's effects on the participating children. If the children progress more than three months, one can infer that the program has had a facilitative effect; however, if the children progress three months or less, one cannot claim the program to have had a facilitative impact. Strategies similar to this are being employed in Title I evaluations (Wisler & Anderson, 1979), although the efficacy of this approach has been recently criticized by Linn (1979).

2) Another analytic approach that can be employed involves the correlation (e.g., Pearson Product Moment) of age at pretest time with pretest score (used by Battelle Corporation; Stock et al., 1976). From this correlation, a predicted posttest score is obtained. The correlated t-test is then run between individual's actual posttest scores and predicted posttest scores. If the differences are significant, with actual posttest scores exceeding predicted scores, the program is presumed to have had an effect. This procedure may assist in determining the educational significance of the obtained differences.

3) A third prediction strategy utilizes the psychometric anchor provided by the Bayley Scales with reference to the infant criterion-referenced measures (e.g., API, SPR). This approach computes the correlation between performance on the Bayley Scales with performance on the other measures. If the correlation is high and positive (e.g., greater than .75), the developmental age scores on the Bayley Scales can be assumed to be an accurate estimator of the developmental age scores of the API and the SPR. Once again, posttest performance on the API and SPR can now be viewed in the context of developmental age progress, a concept which has more meaning to many interventionists and funding agencies.

References

Abt Associates, Education as experimentation: A planned variation model (Vol. 4A). Cambridge, MA: Abt Books, 1977.

Alpern, G., & Boll, T. Developmental profile. Indianapolis, IN: Psychological Development Publications, 1972.

Bayley, N. Manual for the Bayley Scales of Infant Development. New York: Psychological Corporation, 1969.

Bricker, D. Educational synthesizer. In A. Thomas (Ed.), Hey, don't forget about me. Reston, VA: Council for Exceptional Children, 1976.

Bricker, D. Early intervention: The criteria of success. Allied Health and Behavioral Sciences, 1978, 1, 567-582.

Bricker, D. A handicapped children's early educational program: Rationale, program description and impact. Final report for the Division of Innovation and Development, Office of Special Education, January 1981.

Bricker, D., Seibert, J., & Scott, K. Early intervention: History, current status and the problems of evaluation. In D. Doleys, T. Vaughn, & M. Cantrell (Eds.), Interdisciplinary assessment and treatment of developmental problems. New York: Spectrum, Inc., in press.

Bricker, D., & Sheehan, R. Effectiveness of an early intervention program as indexed by measures of child change. Paper presented at HCEEP Project Directors' Meeting, Washington, DC, December 1980.

CAPE, Consortium on Adaptive Performance Evaluation. Adaptive assessment for evaluating the progress of severely/profoundly handicapped children functioning between birth and 2 years. Annual report of a field-initiated research project funded by the Bureau of Education for the Handicapped, Grant No. G007702139, 1978.

Clarke, A., & Clarke, A. Early experience: Myth and evidence. New York: Free Press, 1976.

Cohen, S., Semmes, M., & Guralnick, M. Public Law 94-142 and the education of preschool handicapped children. Exceptional Children, 1979, 45(4), 279-285.

Cook, T., & Campbell, D. Quasi-experimentation design and analysis issues for field settings. Chicago: Rand-McNalley, 1979.

Cross, L., & Johnston, S. A bibliography of instruments. In L. Cross & K. Goin (Eds.), Identifying handicapped children: A guide to casefinding, screening, diagnosis, assessment and evaluation. New York: Walker Publishing Co., 1977.

Fraiberg, S. Parallel and divergent patterns in blind and sighted infants. Psychoanalytic Study of the Child, 1968, 23, 264-300.

Haney, W. A technical history of the national Follow-Through evaluation. Cambridge, MA: Huron Institute, 1977.

Hanson, M. Teaching your Down's syndrome infant: A guide for parents. Baltimore, MD: University Park Press, 1977.

Harbin, G. Mildly to moderately handicapped preschoolers: How do you select child assessment instrument? In T. Black (Ed.), Perspectives on measurement: A collection of readings for educators of young handicapped children. Chapel Hill, NC: Technical Assistance Development System (TADS), 1979.

Helwig, J., & Council, K. (Eds.). SAS user's guide, 1979. Cary, NC, 1979.

Hoepfner, R., Stern, C., & Nummedal, S. CSE-ECRC preschool/kindergarten test evaluations. Los Angeles, CA: School Evaluation Project, Center for the Study of Evaluation and Early Childhood Research Center, Graduate School of Education, University of California, Los Angeles, 1971.

Horst, D., Tallmadge, K., & Wood, C. A practical guide to measuring project impact on student achievement. Washington, DC: U. S. Government Printing Office, 1975.

Johnson, K., & Kopp, C. A bibliography of screening and assessment measures for infants. Project REACH, UCLA Graduate School of Education, Los Angeles, CA, (undated).

Kopp, C., & Shaperman, J. Cognitive development in the absence of object manipulation during infancy. Developmental Psychology, 1973, 9, 430.

Kratochwill, T. Single subject research: Strategies for evaluating change. New York: Academic Press, 1978.

Linn, R. Validity of inferences based on the proposed Title I evaluation models. Educational Evaluation and Policy Analysis, 1979, 1(2), 23-32.

McCarthy, D. Manual for the McCarthy Scales of Children's Abilities. New York: Psychological Corporation, 1972.

Natrella, M. Experimental Statistics: National Bureau of Standards Handbook 91. Washington, DC: U.S. Government Printing Office, 1963.

Nie, N., Hull, C., Jenkins, J., Steinbrenner, K., & Bent, D. Statistical package for the social sciences (2nd ed.). New York: McGraw-Hill, 1975.

Oregon State Mental Health Division. The student progress record, Salem, OR, 1977.

Sheehan, R. Measuring progress in the mild to moderately handicapped preschooler. In T. Black (Ed.), Perspectives on measurement: A collection of readings for educators of young handicapped children. Chapel Hill, NC: Technical Assistance Development System (TADS), 1979.

Sheehan, R. Measuring child progress: Large group design and norm-referenced alternatives. In M. J. May (Ed.), Evaluating handicapped children's early education programs. Proceedings of the 1979 WESTAR Evaluation Workshop, January 31-February 2, 1979, Albuquerque, NM. Seattle, WA: Western States Technical Assistance Resource (WESTAR), 1980.

Stock, J., Wnek, L., Newborg, J., Schenck, E., Gabel, J., Spurgeon, M., & Ray, H. Evaluation of Handicapped Children's Early Education Programs (HCEEP). Columbus, OH: Battelle Center for Improved Education, 1976.

Tallmadge, K. The joint dissemination review panel IDEABOOK. Washington, DC: U.S. Government Printing Office, 1977.

Tallmadge, G., & Horst, D. A procedural guide for validating achievement gains in educational projects. Washington, DC: U.S. Government Printing Office, 1976.

Walls, R., Werner, T., Bacon, A., & Zane, T. Behavior checklists. In J. Cone & R. Hawkins (Eds.), Behavioral assessment: New directions in clinical psychology. New York: Brunner/Mazel, 1977.

White, O., Edgar, E., & Haring, N. Uniform Performance Assessment System. Seattle, WA: Experimental Education Unit, Child Development and Mental Retardation Center, College of Education, University of Washington, 1978.

Wisler, C., & Anderson, J. Designing a Title I evaluation system to meet legislative requirements. Educational evaluation and policy analysis, 1979, 1(2), 47-55.